hydrochloride, ethylene carbonate, thiocyanate, dimethyl sulfoxide and urea. Preferably, the chaotropic auxiliary agent is urea. More preferably, the urea is present at a concentration from about 2.0 to about 8.0 M. Most preferably, the urea is present at a concentration from about 3.0 to about 6.0 M.

In another specific embodiment, the present invention provides a process for obtaining a correctly folded insulin-precursor-containing chimeric protein described above, wherein the incorrectly folded second insulin-precursor-containing chimeric protein is contacted with at least one chaotropic auxiliary agent in an aqueous medium at a pH from about 8.0 to about 10.5 and at a concentration of the incorrectly folded second insulin-precursor-containing chimeric protein from about 0.05 to about 15.0 g per liter. Preferably, the pH is maintained from about 9.0 to about 10.0. Also preferably, the incorrectly folded second insulin-precursor-containing chimeric protein is present from about 0.5 to about 5.0 g per liter. More preferably, the incorrectly folded second insulin-precursor-containing chimeric protein is present from about 2.0 to about 3.0 g per liter.

Mercaptans are compounds which are soluble in water and contain at least one -SH group. In still another specific embodiment, the present invention provides a process for obtaining a correctly folded insulin-precursor-containing chimeric protein described above, further comprising contacting the incorrectly folded second insulin-precursor-containing chimeric protein with a quantity of a mercaptan, which quantity yields less than 5 -SH radical of the mercaptan per cysteine residue of the incorrectly folded second insulin-precursor-containing chimeric protein. In yet another specific embodiment, the incorrectly folded second insulin-precursor-containing chimeric protein is contacted with the mercaptan and the chaotropic auxiliary agent concurrently. In yet another specific embodiment, the incorrectly folded second insulin-precursor-containing chimeric protein is contacted with the mercaptan and the chaotropic auxiliary agent sequentially. Preferably, the quantity of the mercaptan yields from about 0.07 to about 1.0 -SH radical of the mercaptan per cysteine residue of the incorrectly folded second insulin-precursor-containing chimeric protein. Also preferably, the mercaptan is selected from the group consisting of dithiothreitol, dithioerythrol, 2-mercaptoethanol, cysteine, methyl thioglycolate, 3-mercapto-1,2-propanediol and 3-mercaptopropionic acid. More preferably, the mercaptan is 2-mercaptoethanol.

In yet another specific embodiment, the present invention provides a process for obtaining a correctly folded insulin-precursor-containing chimeric protein described above, further comprising separating the correctly folded first insulin-precursor-containing chimeric protein from the incorrectly folded second insulin-precursor-containing chimeric

protein. Preferably, the first insulin-precursor-containing chimeric protein is separated from the second insulin-precursor-containing chimeric protein by ultrafiltration. More preferably, the ultrafiltration is carried out at a pH from about 8.0 to about 11.0. Most preferably, the ultrafiltration is carried out at a pH from about 9.0 to about 10.0.

In yet another specific embodiment, the present invention provides a correctly folded insulin-precursor-containing chimeric protein obtained by the process described above.

4.5. SCREENING AN AMINO ACID SEQUENCE THAT IMPROVES FOLDING OF AN INSULIN PRECURSOR

The present invention provides an assay for screening an amino acid sequence for the ability to improve folding of an insulin precursor, comprising: (a) changing the amino acid sequence of the first peptidyl fragment of a chimeric protein disclosed in Section 4.2., obtaining said chimeric protein with said changes, contacting said chimeric protein with said changes with at least one chaotropic auxiliary agent in an aqueous medium under conditions and for a time sufficient such that said chimeric protein folds correctly, and measuring the folding yield of said chimeric protein with said changes; (b) obtaining the same chimeric protein used in step (a), but without any amino acid sequence changes described in step (a), contacting the chimeric protein without any amino acid sequence changes described in step (a) with the same chaotropic auxiliary agent(s) used in step (a) in an aqueous medium under the same conditions and for a same time used in step (a), and measuring the folding yield of the chimeric protein; and (c) comparing the folding yield of the chimeric proteins measured in step (a) and (b), respectively, in which the yield measured in step (a) substantially equals or is greater than the yield measured in step (b) indicates that the amino acid sequence improves folding of the insulin precursor.

The amino acid sequence of the first peptidyl fragment of a chimeric protein disclosed in Section 4.2. can be changed by any mutagenesis techniques known in the art. preferable by mutagenesis techniques described in Section 4.3.

In a preferred embodiment of the above assay, the chimeric protein consists of the amino acid sequence of SEQ ID NO:6.

In another preferred embodiment of the above assay, the chimeric protein consists of the amino acid sequence of SEQ ID NO:7.

In still another preferred embodiment of the above assay, the chaotropic auxiliary agent is urea.

In yet another preferred embodiment of the above assay, the assay further comprises contacting the chimeric protein, in step (a) and (b) respectively, with a quantity

of a mercaptan, which quantity yields less than 5 -SH radical of the mercaptan per cysteine residue of the chimeric protein. More preferably, the mercaptan is 2-mercaptoethanol.

In a specific embodiment, the product of the above assay is provided.

5. EXAMPLE

A DNA fragment encoding the hGH-mini-proinsulin consisting of the amino acid sequence of SEQ ID NO:6 was chemically synthesized. The DNA fragment was cloned into a bacterial expression vector under the control of a *Trp* promotor. The expression vector containing the hGH-mini-proinsulin was transformed into *E.coli* PR1 strain and the recombinant cells were cultured in M9-CA media in the presence of trace elements. The hGH-mini-proinsulin fusion proteins were recovered from inclusion bodies and folded under the condition such that, within the folded hGH-mini-proinsulin fusion proteins, disulfide bridges were formed as they would be formed in a correctly folded human proinsulin, i.e., the disulfide bridges of A6-A11, A7-B7 and A20-B19 were formed. The folded hGH-mini-proinsulin fusion proteins were separated from the unfolded fusion proteins by a 100K ultrafiltration. The isolated, correctly folded hGH-mini-proinsulin fusion proteins were digested with trypsin to form correctly folded Arg(B31)-human-insulin. The Arg(B31)-human-insulin was purified by cation exchange chromatographies to at least 90% pure. The purified Arg(B31)-human-insulin was then digested with carboxypeptidase B to form the correctly folded human insulin, which was subsequently purified by reversed phase HPLC. The pure human insulin thus produced was characterized by N-terminal sequence analysis, molecular weight determination and peptide mapping.

5.1. CONSTRUCTION OF hGH-MINI-PROINSULIN EXPRESSION VECTOR

A DNA fragment encoding the hGH-mini-proinsulin consisting of the amino acid sequence of SEQ ID NO:6 was chemically synthesized according to the procedure disclosed in Gan et al., *Gene*, 1989, 79:159-166. A 5' Cla I site and a 3' Hind III site were included in the synthesized DNA fragment. Briefly, a fragment from the 5' Cla I to 3' Kpn I, which cuts the nucleotide sequence encoding amino acid residues 51 and 52 of the SEQ ID NO:6, and a fragment from 5' Kpn I to 3 ' Hind III were chemically synthesized and subcloned into a pUC18 vector, respectively. Subsequently, the DNA fragment encoding the entire amino acid sequence of SEQ ID:6 was subcloned into a modified pATH2 vector such that expression of the hGH-mini-proinsulin was under the control of a